Distributed Database Query Management And Storage Techniques For Wireless Sensor Networks

M.Surya[1], S.Muruganandham[2]

M.Phil (Research Scholar), Department of Computer Sciences[1],

Associate Professor, Department of Computer Sciences[2],

Vivekanandha College Of Arts and Sciences For Women(Autonomous) ,Namakkal, India[1,2]

Email:suryaimchandra@gmail.com.

Abstract: Addressing sensor networks, the large amount of data generated by sensors greatly influences the lifetime of the network.To manage this amount of sensed data in an energy-efficient way, new methods of storage and data query are needed. In this way, the distributed database approach for sensor networks is proved as one of the most energy-efficient data storage and query techniques. This paper surveys the state of the art of the techniques used to manage data and queries in wireless sensor networks based on the distributed paradigm. A classification of these techniques is also proposed. The goal of this work is not only to present how data and query management techniques have advanced nowadays, but also show their benefits and drawbacks, and to identify open issues providing guidelines for further contributions in this type of distributed architectures.

KEYWORDS: Distributed database management, wireless sensor networks, distributed storage, query techniques, data reduction techniques, query optimization.

I.INTRODUCTION

Wireless Sensor Networks (WSNs) are composed of a large number of devices, called sensor nodes, which are able to sense, process, and transmit information about the environment on which they are deployed. These devices are usually distributed in a geographical area to collect information for users interested in monitoring and controlling a given phenomenon. This information is transferred to a sink node to be accessible by remote users through generally application-level gateway, e.g. global sensor network (GSN) [1], [2], [3]. To obtain the data, these applications should also provide supports of efficient queries, which allow communication with the network [4], [5], [6] (see Fig. 1 for an illustration of a WSN).In wireless sensor networks, the sensor nodes are battery powered and are considered intelligent with acquisitional, processing, storage, and communication capacities [7], [8]. However, these resources are generally very limited, especially in terms of storage and energy, and the sensor nodes activities are sometimes not negligible in energy consumption [9], [10].

The main goal of distributed database management on WSNs is to support the management of the huge amount of sensed data in an energy-efficient manner [23]. In fact, research into sensor hardware has shown that the energy depletion in the network is mainly due to the data communication tasks among the nodes [24]. To deal with this problem, various data reduction techniques exist [25], [26], including data aggregation [27], [28], packet merging, data compression techniques [29], [30], data fusion, and ap-proximation based techniques [31], [32]. The data aggregation techniques consist to perform data aggregations (e.g., MAX, AVG etc.) at intermediate nodes between the source nodes and the sink node.

The packet merging combines multiple small packets into a big one, without considering the semantics and the correlation between the packets. The data com-pression techniques are also used to reduce the amount of data transmitted between the nodes, but they involve data encoding at the source nodes, data decoding at the sink node.

The data fusion techniques refer to more complex operations on a data set and are usually used in multimedia data processing [33]. The approximation based techniques use statistical techniques to approximate the queries re-sults. These techniques provide, among other advantages, the reduction of the size of the transmitted data, the com-munication tasks, the network load, and the data trans-mission time.

The aim of this paper is to show how distributed database techniques are adapted to wireless sensor net-works to improve the management of the great amount of sensed data in an energy-efficient way by presenting and classifying the most recent and relevant proposals of distributed database management on WSNs. Moreover, a discussion and open issues on distributed database man-agement techniques for wireless sensor networks are identified to facilitate further contributions.

The remainder of this paper is organized as follows. Section 2 presents the essential conceptual features of distributed data storage and querying in WSNs, while protocols and techniques used on the studied proposals on distributed data management in WSNs are exposed in Section 3. Section 4 discusses the techniques used on the studied approaches and proposes some open research issues. Finally, Section 5 concludes the paper and pinpoints further research works.

\

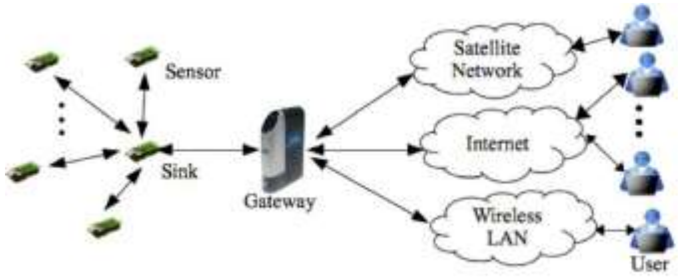

Fig. 1. Illustration of a wireless sensor network architecture

II. RELATED WORK

As largely detailed in there are four essential methods to design a distributed data management system: in-network processing, acquisitional query processing, cross-layer optimization, and data-centric data/query dissemination.

2.1.IN-NETWORK PROCESSING

The in-network processing technique generally includes the different types of operation that are traditionally done on the server, for instance, aggregations to inside the sensor nodes themselves. It is generally used, as its name indicated, to process sensing values inside the network nodes so as to filter and reduce the huge and needless data.

2.2.ACQUISITIONAL QUERY PROCESSING

The acquisitional query processing [20] permits to mini-mize energy consumption in the network by reducing the number of sensor nodes participating in the query processing. This reduction is done by expressing in the query when or what sensors to sample.

2.3.CROSS-LAYER OPTIMIZATION

Unlike the traditional computer networks in which layers in the conventional OSI model are separated and isolated, the cross-layer optimization permits to combine infor-mation available on these different layers and profit from this information sharing. For instance, in wireless sensor net-works the routing takes care of, among others, the quality of service (QoS) parameters of the network, network connec-tivity, the power available on the node, and the network lifetime. In traditional computer networks the routing is done by the network layer only considering the destination address of a packet.

2.4.DATA-CENTRIC DATA/QUERY DISSEMINATION

In contrast with traditional networks, nodes in WSNs usually do not have a single identifier because of data-centric nature of sensor applications as well as the large number of sensors deployed.

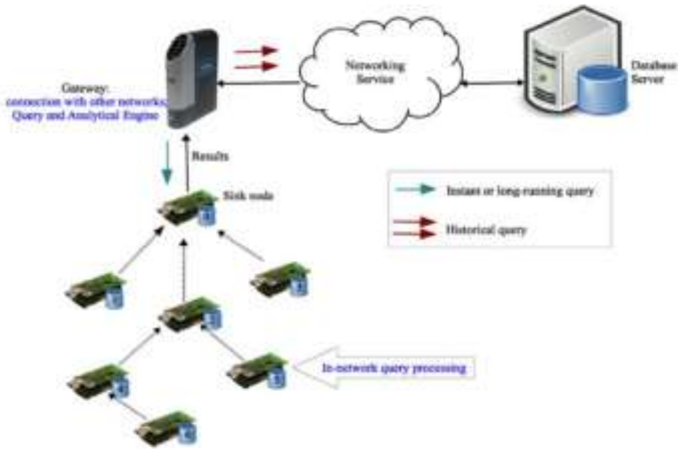

Fig. 2. Illustration of a distributed database architecture for wireless sensor networks

Generally, applications are not interested in specific sensors, but rather in data, which they generate. For example, a query as ‘‘which is the temperature measurement of the sensor with the ID XXXX’’ does not have much interest for a sensor applica-tion, but a query like ‘‘in which region, sensors measure fewer than 7 C’’ is more significant. Routing protocols must take these characteristics into account.

III.DISTRIBUTED DATA STORAGE AND QUERY MANAGEMENT TECHNIQUES FOR WSNS

There is a lot of research addressing various aspects of data management in sensor networks. However, generally, the main goals are the transparent access of the end-users to the sensor nodes for retrieving required relevant sensor data as well as the improvement of the network’s lifetime.

3.1 In-Network Processing

After receiving one query, the query processor analyses it and generates an optimized query execution plan that consists of the best possible execution of the query inside the sensor network. While the different solutions to perform in-network processing may differ, the objective is to save energy by reducing communication. The in-network processing technique can be divided into two sub-categories, aggregation based techniques and approximation based techniques.

3.1.1 Aggregation Based Techniques

Since wireless communication involves more energy consumption than the sensing and processing operations, aggregating the data through intermediate nodes results in less message transmissions. Therefore, the network lifetime is improved with less energy consumption.

As part of the pioneering researches on distributed database on WSNs, TinyDB [20] and COUGAR [21], [34], [49] are the first to adopt a database query optimization technique for WSNs based on the data acquisition declar-ative approach.

3.2.ACQUISITIONAL QUERY PROCESSING

Like in traditional database systems, in distributed data-base systems, where the requested data might be stored in small fragments around the whole network, the complex-ities of mechanisms used by the query processor to efficiently extract the relevant data from the network are completely transparent to the end-users.

In WSNs, the query processor is charged, among other tasks, to generate an optimized query execution plan that defines how a query should be executed in an energy-efficient way. This optimization is performed by minimiz-ing the activities of sensor nodes, principally by reducing the data transmission and sampling the sensors that participate in a query processing.

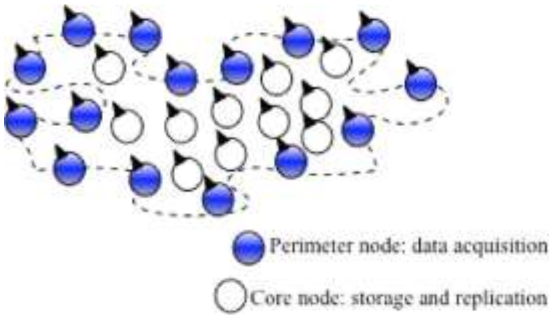

Fig. 3. Framework for the acquisition and storage of spatio-temporal data in mobile sensor networks.

3.3.CROSS-LAYER OPTIMIZATION

The traditional layer approach leads to independent design of different layers and results in strict boundaries between layers. Cross-layer optimization exploits interactions be-tween different layers and can significantly improve energy efficiency as well as adaptability to service, traffic, and environment dynamics. For example, having knowl-edge of the current physical state will help a channel allocation scheme at the MAC layer in optimizing tradeoffs and achieving throughput maximization.
In WSNs, the sensor network lifetime depends intrinsi-cally on the available energy in the nodes composing the network.

This available energy is consumed by the sensing activity, the communication (sending and receiving pack-ets) activity, which is essential to form a WSN, and the data processing [7]. However, the communication activity is more costly in energy than the sensing and processing activities. Hence, current cross-layer optimization techni-ques use a variety of methods to schedule tasks in an energy-efficient way.

The in-network processing is a generic term, which could simply refer to any sort of processing that takes place inside a node. Therefore, the data reduction techniques (data aggregation, packet merging, data compression techniques, data fusion, and approximation based techni-ques) can be placed in the class of in-network processing techniques because each of these techniques is done within nodes composing the network.

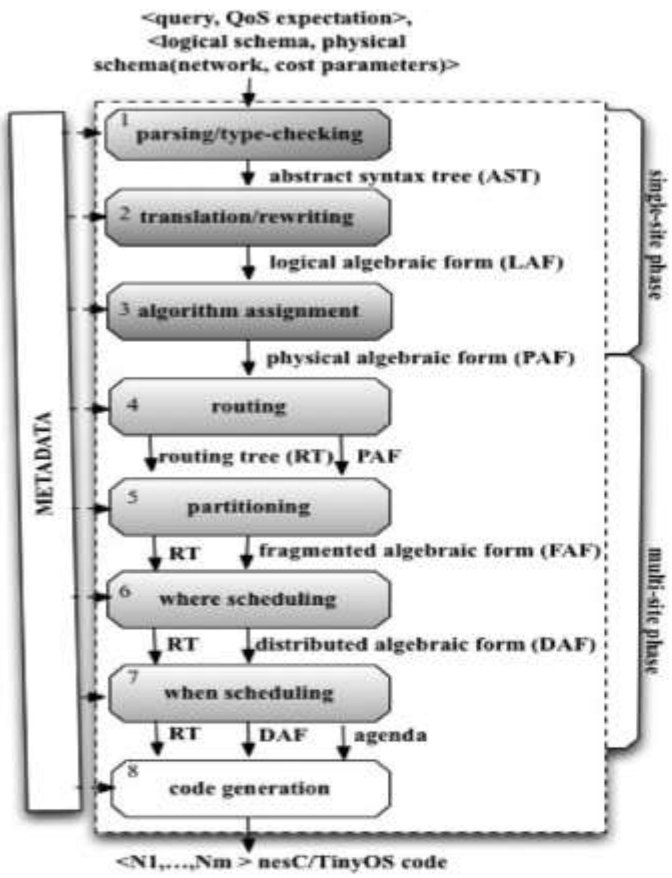

Fig. 4. SNEEql compiler/optimizer architecture.

IV.DISCUSSION AND OPEN ISSUES

The distributed database approach for sensor networks is adopted when sensor nodes do not need to send period-ically the collected data to the base station. The sensed data remains on the sensor nodes and some queries are distributed and processed in the sensor nodes. Thus, in this case, the whole sensor network is viewed as a distributed database. This approach is commonly called in-network processing.

The in-network processing is a generic term, which could simply refer to any sort of processing that takes place inside a node. Therefore, the data reduction techniques (data aggregation, packet merging, data compression techniques, data fusion, and approximation based techni-ques) can be placed in the class of in-network processing techniques because each of these techniques is done within nodes composing the network.

Moreover, one can notice that the three approaches (acquisitional query processing, cross-layer optimization and data-centric data/query dissemination) can be all classed inside the category of in-network processing, since all these three sub-categories require a sensor node to process data and make decision instead of transmitting all the data to a central server for off-line processing, as it is done in in-network processing.

Moreover, these tables highlight the classification of each solution into specific categories.After the detailed analysis of the most recent distributed data management techniques for sensor networks, the following open issues can be identified and suggested:

Keeping and processing sensed data among the sensor nodes are very useful, according to the fact that in-network processing is more energy-efficient than transmitting sensed data for off-line proces-sing. However, this approach may be revised and improved to take into account of the memory overload of the sensor nodes according to the huge amount of generated data compared with the hard resource constraints. Moreover, for efficient query execution in terms of saving energy and good latency, an efficient load balancing policy according to the remaining power and the load of the nodes can be adopted.

. This above study shows that most of the sensor data representations are based on the relational model and the queries are specified using an SQL-like query language. According to the frequent changing characteristics of collected data, the sensor data schema should well fit into XML representation. Then, XPATH can be used for flexible and easy queries.

The distributed technique treats the information within sensor nodes. According to the unstable and generally harsh environment, there may be sudden failures of sensors. This can lead to information loss that greatly influences the result analysis or even the system blocking.

V.CONCLUSION AND FUTURE WORK

Given the great limited sensor device resources, storing and exploiting the large amount of data generated by sensors is a very big problem. The data management techniques used in traditional databases to manage a huge amount of data are not generally suitable for sensor networks taking their specificities into account. Then, the research community has provided a new data storage and querying method, named the distributed database ap-proach for sensor networks. This latter is viewed as the most energy-efficient method to manage the large amount of data generated by the sensor nodes.

Many different techniques have been proposed to manage data and queries in a distributed architecture, but all these techniques can be categorized by their way of processing data into the network (in-network aggregation processing, in-network approximation processing, acquisi-tional query processing, cross-layer optimization, data-centric data/query dissemination). Thus, this work pre-sented and discussed from the oldest to the most recent proposed techniques that have been performed and it is particularly interesting in various stages of distributed data storage, distributed query processing and optimization for sensor networks. The processing techniques aim, generally, to optimize the energy consumption in the network and to retrieve more accurate information. Moreover, the queries used in these techniques are generally SQL-like.

Systems based on sensor networks are increasingly common in many areas of the knowledge, giving rise to several flavors of WSN applications. These applications are generally specific. For instance, there are critical applica-tions that have temporal constraint and need more accurate information to take efficient decisions. New query proces-sing optimization technique, while ensuring data availabil-ity in case of failure is very important in this context. So, one processing technique may not be efficient for the different applications. To deal with these challenges one can opt for a hybrid approach. Furthermore, regarding the Tables 1, 2, and 3, few proposals have been based on metadata management, which is very useful when one manages a huge amount of distributed data. The design of a frame-work that takes into account the various particular char-acteristics of WSN applications can well meet the requirements to a generalized database management sys-tem to handle various types of applications and user needs.

VI.REFERENCES

[1].L. Shu, M. Hauswirth, L. Cheng, J. Ma, V. Reynolds, and L. Zhang, ‘‘Sharing Worldwide Sensor Network,’’ in Proc. IEEE Int. Symp. Appl. Internet, 2008, pp. 189-192.

[2].K. Aberer, M. Hauswirth, and A. Salehi, ‘‘Global Sensor Net-works, Tech. Rep. LSIR-REPORT-2006-001. [Online]. Available: http://lsir.epfl.ch/research/current/gsn/

[3].K. Aberer, M. Hauswirth, and A. Salehi, ‘‘Infrastructure for Data Processing in Large-Scale Interconnected Sensor Networks,’’ in Proc. 8th Int. Conf. Mobile Data Manage., Mannheim, Germany, May 7-11, 2007, pp. 198-205.

[4].E.H. Callaway, Wireless Sensor Networks: Architectures and Proto-cols. Boca Raton, FL, USA: CRC Press, 2004.

[5].P. Baronti, P. Pillai, C.V. Chook, S. Chessa, A. Gotta, and F.Y. Hu, ‘‘Wireless Sensor Networks: A Survey on the State of the Art and the 802.15.4 and ZigBee Standards,’’ J. Comput. Commun., vol. 30, no. 7, pp. 1655-1695, May 2007.

[6].L.M.L. Oliveira and J.J.P.C. Rodrigues, ‘‘Wireless Sensor Net-works: A Survey on Environmental Monitoring,’’ J. Commun., vol. 6, no. 2, pp. 143-151, Apr. 2011.

[7].F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, ‘‘Wireless Sensor Networks: A Survey,’’ Comput. Netw., vol. 38, no. 4, pp. 393-422, Mar. 2002.

[8].A. Nayak and I. Stojmenovic, Wireless Sensor and Actuator Networks: Algorithms and Protocols for Scalable Coordination and Data Communication. Hoboken, NJ, USA: Wiley-Interscience, 2010.

[9].K. Lin, M. Chen, S. Zeadally, and J.J.P.C. Rodrigues, ‘‘Balancing Energy Consumption with Mobile Agents in Wireless Sensor Networks,’’ Future Gener. Comput. Syst., vol. 28, no. 2, pp. 446-456, Feb. 2012.

[10].K. Lin, J.J.P.C. Rodrigues, H. Ge, N. Xiong, and X. Liang, ‘‘Energy Efficiency QoS Assurance Routing in Wireless Multimedia Sensor Networks,’’ IEEE Syst. J., vol. 5, no. 4, pp. 495-505, Dec. 2011.

[11].S. Sendra, J. Lloret, M. Garcia, and J.F. Toledo, ‘‘Power Saving and Energy Optimization Techniques for Wireless Sensor Networks,’’ J. Commun., vol. 6, no. 6, pp. 439-459, Sept. 2011.

[12].S. Matsumae, ‘‘Energy-Efficient Cell Partition of 3D Space for Sensor Networks with Location Information,’’ Netw. Protocols Algorithms, vol. 1, no. 2, pp. 85-98, 2009.

[13].P.A. Neves, J.J.P.C. Rodrigues, M. Chen, and A.V. Vasilakos, ‘‘A Multi-Channel Architecture for IPv6-Enabled Wireless Sensor and Actuator Networks Featuring PnP Support,’’ J. Netw. Comput. Appl., vol. 37, pp. 12-24, Jan. 2014.

[14].L.M.L. Oliveira, J.J.P.C. Rodrigues, A.F. de Sousa, and J. Lloret, ‘‘Network Access Control Framework for 6LoWPAN Net-works,’’ Sensors, vol. 13, no. 1, pp. 1210-1230, Jan. 2013.

[15]..M.L.P. Caldeira, J.J.P.C. Rodrigues, and P. Lorenz, ‘‘Towards Ubiquitous Mobility Solutions for Body Sensor Networks on HealthCare,’’ IEEE Commun. Mag., vol. 50, no. 5, pp. 108-115, May 2012.

[16].A. Campos, E. Souza, F. Nakamura, J.J.P.C. Rodrigues, and E.Nakamura, ‘‘On the Impact of Localization and Density Control Algorithms in Target Tracking Applications for Wire-less Sensor Networks,’’ Sensors, vol. 12, no. 6, pp. 6930-6952, May 2012.

[17].L. Sacks, M. Britton, I. Wokoma, A. Marbini, T. Adebutu, I. Marshall,C.Roadknight, J. Tateson, D. Robinson, and A.G. Velazquez, ‘‘The Development of a Robust, Autonomous Sensor Network Platform for Environmental Monitoring,’’ in Proc. 12th Conf. Sensors Appl., Limerick, Ireland, Sep. 2–4, 2003.

[18].L. Shu, Y. Zhang, G. Min, Y. Wang, and M. Hauswirth, ‘‘Cross-Layer Optimization on Data Gathering in Wireless Multimedia Sensor Networks within Expected Network Lifetime J. Univ. Comput. Sci., vol. 16, no. 10, pp. 1343-1367, 2009.

[19].C. Zhu, L. Shu, T. Hara, L. Wang, S. Nishio, and L.T. Yang, ‘‘A Survey on Communication and Data Management Issues in Mobile Sensor Networks,’’ Wireless Commun. Mobile Comput., vol. 14, no. 1, pp. 19-36, Jan. 2014.

[20].S.R. Madden, M.J. Franklin, J.M. Hellerstein, and W. Hong, ‘‘TinyDB: An Aquisitional Query Processing System for Sensor Networks,’’ ACM Trans. Database Syst., vol. 30, no. 1, pp. 122-173, Mar. 2005.

[21].M. Umer, L. Kulik, and E. Tanin, ‘‘Optimizing Query ProcessingUsing Selectivity-Awarene Wireless Sensor Networks,’’J.Comput., Environ. Urban Syst., vol. 33, no. 2, pp. 79-89, 2009.

[23].L. Cheng, Y. Chen, C. Chen, J. Ma, L. Shu, A.V. Vasilakos, and N.Xiong, ‘‘Efficient Query-Based Data Collection for Mobile Wireless Monitoring Applications,’’ Comput. J., vol. 53, no. 10, pp. 1643-1657, 2010.

[24}..G. Pottie and J.W. Kaiser, ‘‘Wireless Integrated Network Sensors,’’ Commun. ACM, vol. 43, no. 5, pp. 51-58, May 2000.

[25].R. Kacimi, ‘‘Techniques de Conservation d’e´nergie pour les re´seaux de Capteurs Sans Fil,’’ Ph.D. dissertation, Dept. Math., Info. and Telecom., INPT Univ., Toulouse, France, Sept. 2009.

[26].C. Dini, ‘‘Les Re´seaux Capteurs Sans Fil Avec Access Sporadique Au Noeud-puits,’’ Ph.D. dissertation, Info. Eng., Haute Alsace Univ., Mulhouse Cedex, France, Dec. 2010.

[27].S. Madden, M. Franklin, J. Hellerstein, and W. Hong, ‘‘TAG: A Tiny Aggregation Service for Ad-Hoc Sensor Networks,’’ in Proc. ACM Symp. OSDI, Boston, MA, USA, 2002, pp. 131-146.

[28].E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, ‘‘In-Network Aggregation Techniques for Wireless Sensor Networks: A Sur-vey,’’ IEEE Wireless Commun., vol. 14, no. 2, pp. 70-87, Apr. 2007.

[29].C. Tang and C.S. Raghavendra, ‘‘Compression Techniques for Wireless Sensor Networks Wireless Sensor Networks. Norwell, MA, USA: Kluwer, 2004, pp. 207-231.

[30].C. Dini and P. Lorenz, ‘‘Primitive Operations for Prioritized Data Reduction in Wireless Sensor Network Nodes,’’ in Proc. 4th ICSNC, Porto, Portugal, 2009, pp. 274-280.

[31]. Xu, X. Tang, and W.C. Lee, ‘‘A New Storage Scheme for Ap-proximate Location Queries in Object Tracking Sensor Networks,’’ IEEE Trans. Parallel Distrib. Syst., vol. 19, no. 2, pp. 262-275, Feb. 2008.

